

Apriori Algorithm through RapidMiner for Age Patterns of Homeless and Beggars

¹Wirta Agustin, ²Yulya Muharmi

¹Jurusan Sistem Informasi, STMIK Amik Riau

²Jurusan Sistem Informasi, STMIK Dharmapala Riau

Email: ¹wirtaagustin@stmik-amik-riau.ac.id, ²ymuharmi@gmail.com

Article Info

Article history:

Received Aug 15th, 2018

Revised Aug 20th, 2018

Accepted Sept 26th, 2018

Keyword:

Age

Apriori Algorithm

Data Mining

Homelessness and Beggars

RapidMiner

ABSTRACT

Homeless and beggars are one of the problems in urban areas as they possibly disrupt public order, security, stability and urban development. The efforts conducted are still focusing on managing the existing homeless and beggars instead of preventing the potential ones. One of the methods used for solving this problem is Algoritma Apriori which determines the age pattern of homeless and beggars. Apriori Algorithm is an Association Rule method in data mining to determine frequent item set that serves to help in finding patterns in a data (frequent pattern mining). The manual calculation through Apriori Algorithm obtains combination pattern of 11 rules with a minimum support value of 25% and the highest confidence value of 100%. These patterns were references for the incharged department in precaution action of homeless and beggars arising numbers. Apriori Algorithm testing uses the RapidMiner application which is one of data mining processing software, including text analysis, extracting patterns from data sets and combining them with statistical methods, artificial intelligence, and databases to obtain high quality information from processed data. Based on the results of the said testing, it can be concluded that the level of accuracy test is low, i.e. 37.5%.

Copyright © 2018 Puzzle Research of Data Technology

Corresponding Author:

Wirta Agustin,

Program Studi Sistem Informasi,

STMIK Amik Riau,

Jl. Purwodadi Indah Km. 10 Panam, Kel. Sialangmunggu, Kec. Tampan, Kota Pekanbaru, 28294, Riau.

Email: wirtaagustin@stmik-amik-riau.ac.id

1. PENDAHULUAN

Gelandangan dan pengemis (GePeng) salah satu masalah yang ada di daerah perkotaan. Masalah umum gelandangan dan pengemis berkaitan dengan masalah kebersihan, ketertiban dan keamanan di daerah perkotaan, yang akan mengganggu stabilitas dan pembangunan kota. Upaya penanganan gelandangan dan pengemis sudah dilakukan secara intens, seperti penertiban, pembinaan dan pemberdayaan masyarakat. Semua upaya yang dilakukan selama ini masih fokus pada bagaimana cara penanganan gelandangan dan pengemis yang sudah ada, belum dilakukan upaya untuk pencegahan gelandangan dan pengemis.

Penelitian terkait gelandangan dan pengemis sebelumnya sudah pernah dilakukan, tetapi baru dalam bidang ilmu sosiologi dan ilmu hukum, diantaranya: “Faktor Penyebab Terjadinya Gelandangan Dan Pengemis (Studi Pada Gelandangan Dan Pengemis Di Kecamatan Tanjung Karang Pusat Kota Bandar Lampung)” [1], “Penanggulangan Pengemis Dan Gelandangan Di Kota Pekanbaru” [2], dan “Strategi Pelaksanaan Pembinaan Gelandangan dan Pengemis oleh Dinas Sosial dan Pemakaman Kota Pekanbaru” [3].

Pentingnya upaya pencegahan gelandangan dan pengemis serta merujuk pada penelitian yang telah pernah dilakukan sebelumnya, maka perlu dilakukan penelitian dalam bidang informatika, khususnya data mining untuk mengetahui pola penyebab gelandangan dan pengemis dari database yang ada. Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma dalam data mining sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses Karakteristik Data Mining (KDD) secara keseluruhan [4].

Salah satu upaya yang bisa dilakukan adalah dengan menentukan pola usia gelandangan dan pengemis. Algoritma Apriori sebuah metode *Association Rule* dalam data mining untuk menentukan *frequent itemset* yang berfungsi membantu menemukan pola dalam sebuah data (*frequent pattern mining*) [5].

Pengujian penerapan Algoritma Apriori menggunakan aplikasi RapidMiner. RapidMiner salah satu *software* pengolahan data mining, diantaranya analisis teks, mengekstrak pola-pola dari data set dan mengkombinasikannya dengan metode statistika, kecerdasan buatan, dan database untuk mendapatkan informasi bermutu tinggi dari data yang diolah [4]. Hasil pengujian menunjukkan perbandingan pola usia gelandangan dan pengemis yang berpotensi menjadi gelandangan dan pengemis dari pengujian dengan aplikasi RapidMiner dan perhitungan manual menggunakan Algoritma Apriori.

2. METODOLOGI PENELITIAN

Pengumpulan data dilakukan melalui wawancara langsung dengan bagian terkait Dinas Sosial dan Pemakaman Kota Pekanbaru dan Satpol PP Pekanbaru, serta mengumpulkan data gelandangan dan pengemis periode tahunan. Data yang diolah sebanyak 250 data gelandangan dan pengemis berdasarkan kategori usia Balita (0 - 5 tahun), Kanak-kanak (5 - 11 tahun), Remaja Awal (12 - 16 tahun), Remaja Akhir (17 - 25 tahun), Dewasa Awal (26- 35 tahun), Dewasa Akhir (36- 45 tahun), Lansia Awal (46- 55 tahun) dan Lansia Akhir (56 - 65 tahun) [6].

Pengujian penerapan Algoritma Apriori menggunakan aplikasi RapidMiner untuk pola usia gelandangan dan pengemis, dilakukan dalam 4 tahapan, yaitu: (1) skenario pengujian, (2) kriteria pengujian, (3) proses pengujian, dan (4) hasil pengujian.

2.1. Skenario Pengujian

Dalam melakukan pengujian pola usia gelandangan dan pengemis, dengan penerapan Algoritma Apriori menggunakan aplikasi RapidMiner akan dilakukan tahapan sebagai berikut : (1) membuat tabular dalam format data ms. *excel* yang berisi tentang data-data usia gelandangan dan pengemis yang sudah dikelompokkan (lihat 2.1.1); (2) menentukan kriteria pengujian data-data usia gelandangan dan pengemis yang akan digunakan sebagai acuan atau tolak ukur dalam melihat hasil pengujian, apakah pengujian yang dilakukan memiliki tingkat keakuratan pengujian sangat tinggi, tinggi dan rendah (lihat 2.2); (3) melakukan proses pengujian data-data usia gelandangan dan pengemis menggunakan aplikasi RapidMiner (lihat 2.3); dan (4) hasil pengujian data-data usia gelandangan dan pengemis yang didapat dari hasil aplikasi RapidMiner (lihat 2.4).

2.1.1. Tabular Data Usia Gelandangan dan Pengemis dalam Format Ms. Excel

Data gelandangan dan pengemis berdasarkan kategori usia dibuat tabel tabular dalam format ms. *excel* yang dapat mempermudah dalam mengetahui berapa banyak item yang ada dalam setiap bulan seperti pada tabel 1[7].

Tabel 1. Format Tabular Data Usia Gelandangan dan Pengemis

Bulan	Balita	Kanak	Remaja Awal	Remaja Akhir	Dewasa Awal	Dewasa Akhir	Lansia Awal	Lansia Akhir
1	0	0	0	1	0	1	0	1
2	0	0	0	1	1	1	0	0
3	0	0	0	1	1	1	0	0
4	0	0	1	1	1	0	0	0
5	0	0	0	1	0	1	1	0
6	0	1	0	1	0	0	1	0
7	0	0	1	1	1	0	0	0
8	0	0	1	1	0	0	0	1
9	0	0	0	1	1	1	0	0
10	0	0	0	1	1	1	0	0
11	0	0	1	1	1	0	0	0
12	0	0	1	0	1	1	0	0

2.2. Kriteria Pengujian

Kriteria yang digunakan dalam pengujian data-data usia gelandangan dan pengemis adalah berdasarkan pada pola (*rules*) dan nilai *confidence* (c) yang diperoleh dari hasil pengujian aplikasi RapidMiner, yaitu: pola (*rules*) sebanyak 2(dua) pola (*rules*) dan nilai *confidence* (c) = 80%. (1) jika pola (*rules*) dan nilai *confidence* (c) hasil perhitungan manual sama dengan nilai hasil pengujian aplikasi RapidMiner, maka tingkat keakuratan pengujian sangat tinggi; (2) jika pola (*rules*) dan nilai *confidence* (c) hasil perhitungan manual mendekati nilai hasil pengujian aplikasi RapidMiner, maka tingkat keakuratan pengujian tinggi; dan (3) jika pola (*rules*) dan nilai *confidence* (c) hasil perhitungan manual tidak mendekati nilai hasil pengujian aplikasi RapidMiner, maka tingkat keakuratan pengujian rendah.

2.3. Proses Pengujian

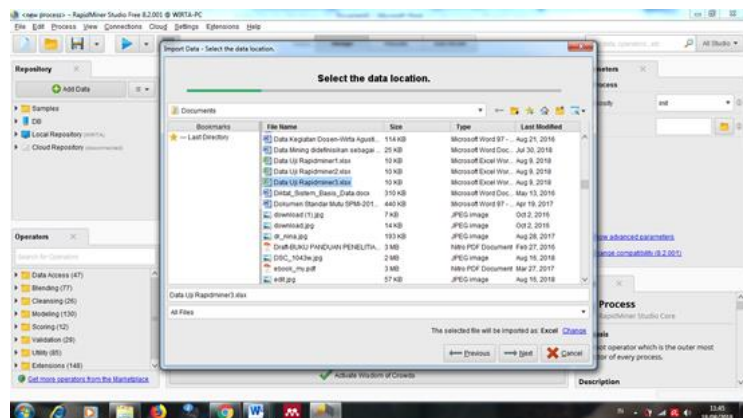
Aplikasi yang digunakan dalam pengujian ini adalah RapidMiner Studio version 8.2. Proses pengujian dengan aplikasi RapidMiner terdiri dari beberapa tahap, yaitu[8]: (1) add data adalah proses pengambilan data tabular usia gelandangan dan pengemis yang telah disimpan dalam format data excel (lihat 2.3.1); (2) memilih data, proses memilih data dilakukan untuk melihat dan memeriksa data yang akan digunakan (lihat 2.3.2); (3) proses Algoritma Apriori adalah proses menggabungkan data dengan operasi-operasi terkait Algoritma Apriori (lihat 2.3.3); (4) hasil running Algoritma Apriori, proses melihat hasil dari penerapan Algoritma Apriori pada aplikasi RapidMiner (lihat 2.3.4).

2.3.1. Add Data

Proses pengambilan data tabular usia gelandangan dan pengemis yang telah disimpan dalam format data menggunakan ms. Excel. Add data dilakukan dengan cara mengklik menu Add Data, muncul tempat data, kemudian pilih file yang akan ditambahkan, dapat dilihat pada gambar 1 dan gambar 2.



Gambar 1. Add Data dari Ms. Excel



Gambar 2. Pilih File

2.3.2. Memilih Data

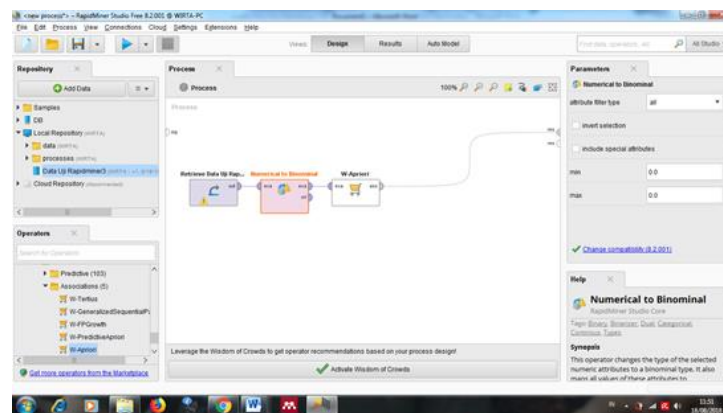
Proses memilih data dilakukan untuk melihat dan memeriksa data yang akan digunakan. Data yang dipilih hanya kolom usia saja tidak termasuk bulan, dapat dilihat pada gambar 3.

Row No.	BALITA	KANAK-KAN.	REMAJA AW.	REMAJA AK.	DEWASA A.	DEWASA AK.	LANSIA AW.	LANSIA AK.
1	0	0	0	1	0	1	0	1
2	0	0	0	1	1	1	0	0
3	0	0	0	1	1	1	0	0
4	0	0	1	1	1	0	0	0
5	0	0	0	1	0	1	1	0
6	0	1	0	1	0	0	1	0
7	0	0	1	1	1	0	0	0
8	0	0	1	1	0	0	0	1
9	0	0	0	1	1	1	0	0
10	0	0	0	1	1	1	0	0
11	0	0	1	1	1	0	0	0
12	0	0	1	0	1	1	0	0

Gambar 3. Memilih Data

2.3.3. Proses Algoritma Apriori

Setelah melakukan proses add data dan memilih data, maka akan dilakukan proses Algoritma Apriori dengan cara drag and drop pada tab repository dan operators[9]. Proses menggabungkan data tabular dengan operator-operator terkait Algoritma Apriori, dapat dilihat pada gambar 4.



Gambar 4. Proses Algoritma Apriori

2.3.4. Hasil Running Algoritma Apriori

Hasil *running* Algoritma Apriori merupakan hasil data berupa *rules*. Hasil *running* dapat dilihat dengan mengklik tab *results* atau F9 pada *keyboard*, maka akan tampil 3 tab (*Result History*, *W-Apriori*, *ExampleSet*). Tab *Result History* adalah tab yang menampilkan *history* proses Algoritma Apriori, tab *W-Apriori* adalah tab yang menampilkan hasil algoritma apriori berupa *rules*, dan tab *ExampleSet* menampilkan data tabular yang digunakan[10]. Proses ini dapat dilihat pada gambar 5.

W-Apriori

Minimum support: 0.9 (9 instances)
 Minimum interest confidence: 0.4
 Number of cycles performed: 14

Generated sets of large itemsets:

Size of set of large itemsets S(1): 2

Large itemsets S(1):
 REMAJA AKER = true & DEWASA AKER = false & DEWASA AKER = false

Size of set of large itemsets S(2): 1

Large itemsets S(2):
 REMAJA AKER = true & REMAJA AKER = false & DEWASA AKER = false

Best rules found:

1. DEWASA AKER = false & REMAJA AKER = true & DEWASA AKER = false (0.9)
 2. REMAJA AKER = true & REMAJA AKER = false & DEWASA AKER = false (0.9)

Gambar 5. Hasil Running Algoritma Apriori

2.4. Hasil Pengujian

Proses pengujian dengan aplikasi RapidMiner, menemukan 2 (dua) pola (*rules*) terbaik dan nilai *confidence* (c) sebesar 80%, (lihat gambar 5). Hasil pengujian dapat dilihat pada tabel 2.

Tabel 2. Hasil Pengujian Aplikasi RapidMiner

Aturan	Confidence
Jika Gepeng Kategori Usia Dewasa Akhir, Maka Gepeng Juga Kategori Usia Remaja Awal	80%
Jika Gepeng Kategori Usia Remaja Awal, Maka Gepeng Juga Kategori Usia Dewasa Akhir	80%

Dari tabel 2 Dapat diambil kesimpulan bahwa usia gelandangan dan pengemis yang berpotensi menjadi gelandangan dan pengemis adalah kategori usia dewasa akhir (36 s/d 45 tahun) dan remaja awal (12 s/d 16 tahun).

3. HASIL DAN ANALISIS

Hasil pengujian Algoritma Apriori menggunakan aplikasi RapidMiner dibandingkan dengan hasil perhitungan manual menggunakan Algoritma Apriori yang telah dilakukan, kemudian dilakukan analisis sehingga dapat disimpulkan tingkat keakuratan pengujian sesuai dengan kriteria pengujian yang telah ditentukan, apakah pola (*rules*) dan nilai *confidence* (c) hasil perhitungan manual Algoritma Apriori sama, mendekati atau tidak mendekati pola (*rules*) dan nilai *confidence* (c) dari hasil pengujian aplikasi RapidMiner. Proses pengujian dengan aplikasi RapidMiner, menemukan 2 (dua) pola (*rules*) terbaik dan nilai *confidence* (c) sebesar 80%, (lihat gambar 5 dan tabel 2). Hasil perhitungan manual menggunakan Algoritma Apriori, menemukan 3 (tiga) pola (*rules*) terbaik dan nilai *confidence* (c) bervariasi sebesar 63%, 71% dan 100%, dapat dilihat tabel 3.

Tabel 3. Hasil Perhitungan Manual Menggunakan Algoritma Apriori

Aturan	Confidence
Jika Gepeng Kategori Umur Remaja Akhir, Dewasa Awal, Maka Gepeng Juga Kategori Umur Dewasa Akhir	63%
Jika Gepeng Kategori Umur Remaja Akhir, Dewasa Akhir, Maka Gepeng Juga Kategori Umur Dewasa Awal	71%
Jika Gepeng Kategori Umur Dewasa Awal, Dewasa Akhir, Maka Gepeng Juga Kategori Umur Remaja Akhir	100%

Dari tabel 3. Dapat diambil kesimpulan bahwa usia gelandangan dan pengemis yang berpotensi menjadi gelandangan dan pengemis adalah kategori usia remaja akhir (17 s/d 25 tahun), dewasa awal (26 s/d 35 tahun) dan dewasa akhir (36 s/d 45 tahun). Berikut adalah perbandingan hasil yang diperoleh dari hasil pengujian dengan aplikasi RapidMiner dan perhitungan manual menggunakan Algoritma Apriori (lihat tabel 2 dan tabel 3):

Tabel 4. Perbandingan Hasil Pengujian Algoritma Apriori dan Aplikasi RapidMiner

Pengujian	Kategori Usia	Confidance	Rules
Algoritma Apriori	Remaja Akhir, Dewasa Awal, Dewasa Akhir	63%, 71%, 100%	3 (tiga)
RapidMiner	Remaja Awal, Dewasa Akhir	80%, 80%	2 (dua)
Selisih	Remaja Akhir, Dewasa Awal, Remaja Awal	17%, 9%, 20%	1 (satu)

4. KESIMPULAN

Berdasarkan hasil pengujian aplikasi RapidMiner dan hasil perhitungan manual Algoritma Apriori yang telah dilakukan, dapat disimpulkan sesuai dengan kriteria pengujian yang telah ditentukan, bahwa pola (*rules*) usia dan nilai *confidence* (c) hasil perhitungan manual Algoritma Apriori tidak mendekati nilai hasil pengujian menggunakan aplikasi RapidMiner, maka tingkat keakuratan pengujian rendah. Berdasarkan hasil penelitian yang penulis lakukan dan hasil penelitian sebelumnya dapat disimpulkan perbedaan yang signifikan, yaitu hasil penelitian sebelumnya berupa faktor internal (kemiskinan, keluarga, umur, cacat fisik, rendahnya pendidikan, rendahnya keterampilan, dan sikap mental) dan faktor eksternal (lingkungan, letak geografis dan lemahnya penanganan masalah gelandangan dan pengemis) penyebab terjadinya gelandangan dan pengemis. Hambatan dalam penertiban dan pembinaan gelandangan dan pengemis. Strategi pelaksanaan pembinaan gelandangan dan pengemis menggunakan indikator Perencanaan, Pemosisian dan Penilaian. Sedangkan hasil dari penelitian yang penulis lakukan berupa pola penyebab gelandangan dan pengemis

berdasarkan kategori usia Balita, Kanak-kanak, Remaja Awal, Remaja Akhir, Dewasa Awal, Dewasa Akhir, Lansia Awal dan Lansia Akhir.

UCAPAN TERIMA KASIH

Ucapan terimakasih kepada Kemenristek Dikti untuk pendanaan hibah Penelitian Dosen Pemula tahun pendanaan 2018, LPPM STMIK Amik Riau, Dinas Sosial dan Pemakaman Kota Pekanbaru, Satpol PP Kota Pekanbaru dan semua pihak yang telah membantu dalam melakukan penelitian ini.

REFERENSI

- [1] I. Riskawati and A. Syani, "FAKTOR PENYEBAB TERJADINYA GELANDANGAN DAN PENGEMIS (Studi Pada Gelandangan dan Pengemis Di Kecamatan Tanjung Karang Pusat Kota Bandar Lampung)," *J. Sociol.*, vol. 1, no. 1, pp. 43–52.
- [2] M. S. Muslim, S.Sos, "Penanggulangan Pengemis Dan Gelandangan Di Kota Pekanbaru," no. 40, 1983.
- [3] S. Pelaksanaan, P. Gelandangan, and K. Pekanbaru, "Strategi Pelaksanaan Pembinaan Gelandangan dan Pengemis oleh Dinas Sosial dan Pemakaman Kota Pekanbaru."
- [4] R. Yanto and R. Khoiriah, "Implementasi Data Mining dengan Metode Algoritma Apriori dalam Menentukan Pola Pembelian Obat," *citec J.*, vol. 2, pp. 102–113, 2015.
- [5] R. Sistem and E. Elisa, "JURNAL RESTI Dengan Algoritma Apriori," *RESTI*, vol. 2, no. 2, pp. 472–478, 2018.
- [6] "Kategori Umur Depkes RI," 2009. [Online]. Available: <https://www.scribd.com/doc/151484440/Kategori-Umur-Menurut-Depkes-RI>.
- [7] P. Data, Y. Digunakan, D. Minggu, D. I. Bulan, and O. Tahun, "ALGORITMA APRIORI UNTUK REKOMENDASI PENAWARAN PRODUK DI," vol. 8, 2015.
- [8] Shalsabilla Luthfi Dewati, "ANALISA POLA TRANSAKSI OBAT MENGGUNAKAN ALGORITMA APRIORI," 2014.
- [9] U. Mazida, R. A. Pramunendar, F. I. Komputer, and U. D. Nuswantoro, "ANALISIS ALGORITMA APRIORI UNTUK REKOMENDASI PENEMPATAN BUKU PADA PERPUSTAKAAN," pp. 1–8.
- [10] B. R. C. T. I, A. A. Gafar, N. Fajriani, U. Ramdani, F. R. Uyun, Y. P. P, and N. Ransi, "Implementasi k-means clustering pada rapidminer untuk analisis daerah rawan kecelakaan," no. April, pp. 58–62, 2017.
- [11] A. Primajaya, B. N. Sari, T. Informatika, F. I. Komputer, U. S. Karawang, A. Info, and R. Forest, "Random Forest Algorithm for Prediction of Precipitation," *Indones. J. Artif. Intell. Data Min.*, vol. 1, no. 1, pp. 27–31, 2018.

BIBLIOGRAFI PENULIS



Wirta Agustin, M.Kom, lahir di Solok pada 10 Agustus 1972. Saat ini penulis adalah dosen pada Program Studi Sistem Informasi, STMIK Amik Riau. Lulus S1 dari Program Studi Manajemen Informatika, STMIK "YPTK" Padang pada tahun 1998, lalu melanjutkan studi S2 pada Program Studi Magister Ilmu Komputer, Konsentrasi Sistem Informasi pada Universitas Putra Indonesia "YPTK" Padang, lulus pada tahun 2013.



Yulya Muharmi, M.Kom, lahir di Bukittinggi pada 02 Bukittinggi 1992. Lulus S1 dari program studi Sistem Informasi, Universitas Putra Indonesia "YPTK" Padang tahun 2014. Pada tahun 2014, penulis melanjutkan studi S2 di Universitas Putra Indonesia "YPTK" Padang dengan program studi Magister Ilmu Komputer dengan Konsentrasi Sistem Informasi dan lulus pada tahun 2015. Saat ini penulis adalah dosen di Program Studi Sistem Informasi di STMIK Dharmapala Riau.